

AVERAGING ALMOST PERIODIC FUNCTIONS ALONG EXPONENTIAL SEQUENCES

MICHAEL BAAKE, ALAN HAYNES, AND DANIEL LENZ

ABSTRACT. The goal of this expository article is a fairly self-contained account of some averaging processes of functions along sequences of the form $(\alpha^n x)_{n \in \mathbb{N}}$, where α is a fixed real number with $|\alpha| > 1$ and $x \in \mathbb{R}$ is arbitrary. Such sequences appear in a multitude of situations including the spectral theory of inflation systems in aperiodic order. Due to the connection with uniform distribution theory, the results will mostly be *metric* in nature, which means that they hold for Lebesgue-almost every $x \in \mathbb{R}$.

1. INTRODUCTION

A frequently encountered problem in mathematics and its applications is the study of averages of the form $\frac{1}{N} \sum_{n=1}^N f(x_n)$, where f is a function with values in \mathbb{C} or, more generally, in some Banach space, and $(x_n)_{n \in \mathbb{N}}$ is a sequence of numbers in the domain of f . Quite often, an exact treatment of these averages is out of hand, and one resorts to the analysis of asymptotic properties for large N . This, for instance, is common in analytic number theory; compare [19, 20, 1] and references therein. Equally important is the case where one can establish the existence of a limit as $N \rightarrow \infty$, and then calculate it. This occupies a good deal of ergodic theory, where Birkhoff's theorem and Kingman's subadditive theorem provide powerful tools to tackle the problem; see [15, 38] for background.

However, not all tractable cases present themselves in a way that is immediately accessible to tools from ergodic theory. Also, depending on the nature of the underlying problem, one might prefer a more elementary method, as Birkhoff-type theorems already represent a fairly advanced kind of 'weaponry'. An interesting (and certainly not completely independent) approach is provided by the theory of uniform distribution of sequences, which essentially goes back to Weyl [39] and has emerged as a major tool for the study of function averages, in particular for functions that are periodic or defined on a compact domain; see [25, 17, 26] and references therein for more.

In this contribution, we recall some of these concepts, with an eye on both methods (uniform distribution and ergodic theory), and use the tools to treat averages of almost periodic functions along sequences where this makes sense, in particular along sequences of the form $(\alpha^n x)_{n \in \mathbb{N}}$ with 'generic' $x \in \mathbb{R}$ and a fixed number $\alpha \in \mathbb{R}$ with $|\alpha| > 1$. The first subtlety that we shall encounter here emerges when α is not an integer, which requires some care for functions that fail to be locally Riemann-integrable. The second subtlety occurs when we extend our considerations to almost periodic functions.

While the latter extension represents a relatively simple step beyond periodic functions as long as one retains almost periodicity in the sense of Bohr, matters become more involved when singularities occur or weaker notions of almost periodicity are needed. Below, we shall discuss some extensions of this kind that are relevant in practice; compare [18] for some related results. Let us note that some of the notions and concepts used below are studied in much greater generality in [29, 36].

Before we begin our exposition, let us mention that averages of 1-periodic functions are often just the first step in the study of Riesz–Raikov sums, that is, sums of the form $\sum_{k=0}^{n-1} f(\alpha^k t)$. Kac’s investigation for $\alpha = 2$ in [22] and Takahashi’s refined and generalised analysis [37] are early examples that consider limits (in a law of large numbers scaling) as well as distributions (in a central limit theorem scaling, when $\int_0^1 f(t) dt = 0$). This led to a more elaborate derivation of central limit theorems for Riesz–Raikov sums along exponential sequences; compare [31, 27, 33] and references therein.

Below, we are mainly interested in the Birkhoff-type averages, with a focus on functions that fail to be periodic, but still have some repetitivity structure in the form of a suitable almost periodicity. In this sense, we have selected one particular aspect of Riesz–Raikov sums that appears in the theory of aperiodic order [3, 10, 2].

2. PRELIMINARIES AND GENERAL SETTING

As far as possible, we follow the general (and fairly standard) notation from [3, Ch. 1], wherefore only deviations or extensions will be mentioned explicitly. In particular, we will use the Landau symbols \mathcal{O} and \mathcal{o} for the standard asymptotic behaviour of real- or complex-valued functions; compare [1, 19] for definitions and examples.

When two sets $A, B \subseteq \mathbb{R}$ are given, we denote their *Minkowski sum* as

$$A + B := \{a + b : a \in A, b \in B\}.$$

In particular, if the point set $S \subset \mathbb{R}$ is locally finite and $\varepsilon > 0$, we use $S + (-\varepsilon, \varepsilon)$ for the open subset of \mathbb{R} that emerges from S as $\bigcup_{x \in S} (x - \varepsilon, x + \varepsilon)$. Note that its complement in \mathbb{R} is then a closed set (possibly empty).

Below, we frequently talk about results of metric nature, where Lebesgue measure λ on \mathbb{R} is our reference measure. When a statement is true for almost every $x \in \mathbb{R}$ with respect to Lebesgue measure, we will simply say that it holds for a.e. $x \in \mathbb{R}$. Likewise, when we speak of a null set, we mean a null set with respect to Lebesgue measure.

Recall that a sequence $(x_n)_{n \in \mathbb{N}}$ of real numbers is called *uniformly distributed modulo 1* if, for all real numbers a, b with $0 \leq a < b \leq 1$, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \text{card}\left([a, b) \cap \{\langle x_1 \rangle, \dots, \langle x_N \rangle\}\right) = b - a,$$

where $\langle x \rangle$ denotes the fractional part¹ of $x \in \mathbb{R}$. We refer to [25, 11] for general background. Recall that a function f on \mathbb{R} is 1-periodic if $f(x+1) = f(x)$ holds for all $x \in \mathbb{R}$. One fundamental result, due to Weyl [39], can now be formulated as follows; see also [20, Thm. 5.3].

Lemma 2.1 (Weyl's criterion). *For a sequence $(x_n)_{n \in \mathbb{N}}$ of real numbers, the following properties are equivalent.*

- (1) *The sequence is uniformly distributed modulo 1.*
- (2) *For every complex-valued, 1-periodic continuous function f , one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_0^1 f(x) dx.$$

- (3) *The relation from (2) holds for every 1-periodic function that is locally Riemann-integrable.*
- (4) *The relation*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i k x_n} = \delta_{k,0}$$

holds for every $k \in \mathbb{Z}$. □

Let us note in passing that the equivalence of conditions (1) and (2) can also be understood in terms of systems of almost invariant integrals and Eberlein's ergodic theorem. These notions are reviewed and studied in some detail in [36].

Remark 2.2. Weyl's criterion (which is also known as Weyl's lemma) is an important tool for calculating the average of a locally Riemann-integrable periodic function along a uniformly distributed sequence. In fact, a 1-periodic function is locally Riemann-integrable if and only if the Birkhoff average converges for *every* sequence that is uniformly distributed modulo 1; compare [16] as well as [20, p. 123].

Conversely, the integral of a Riemann-integrable function can be approximated by averages along uniformly distributed sequences. This is a standard method in numerical integration, in particular for higher-dimensional integrals; see [21, 26] and references therein for more. \diamond

There is an abundance of known results on uniformly distributed sequences and their finer properties; we refer to [25] for the classic theory and to [11] and references therein for more recent developments. Here, we are particularly interested in one specific class of sequences, for which the uniform distribution is well known; compare [11, Thms. 1.7 and 1.10] as well as [15, Sec. 7.3, Thm. 1] or [25, Cor. 1.4.3 and Exs. 1.4.3].

Fact 2.3. *Consider the sequence $(\alpha^n x)_{n \in \mathbb{N}}$. For fixed $\alpha \in \mathbb{R}$ with $|\alpha| > 1$, it is uniformly distributed modulo 1 for a.e. $x \in \mathbb{R}$. For fixed $0 \neq x \in \mathbb{R}$, the sequence is uniformly distributed modulo 1 for a.e. $\alpha \in \mathbb{R}$ with $|\alpha| > 1$.* □

¹Since we use $\{x\}$ for singleton sets, we resort to the less common notation $\langle x \rangle$ for the fractional part of x in order to avoid misunderstandings.

Below, we will mainly be concerned with the first case, where a fixed α with $\alpha > 1$ or $|\alpha| > 1$ is given. This situation is of particular interest in the theory of aperiodic order, for instance in connection with the renormalisation analysis of inflation tiling systems, because it plays an important role for the averaging of functions with certain repetition properties along the real line.

Let us mention in passing that, when $\alpha = q \geq 2$ is an integer, $(q^n x)_{n \in \mathbb{N}}$ is uniformly distributed modulo 1 if and only if x is a *normal* number [11] in base q , which means that the q -ary expansion of x contains all possible finite substrings in the digit set $\{0, 1, \dots, q-1\}$ in such a way that any substring of length ℓ has frequency $1/q^\ell$. In the Lebesgue sense, a.e. $x \in \mathbb{R}$ is normal with respect to all integer bases [11, Thm. 4.8], but it is a hard problem to decide on normality for any given number.

Remark 2.4. Consider a sequence $(u_n)_{n \in \mathbb{N}_0}$ of real numbers such that $\inf_{n \neq m} |u_n - u_m| > 0$. Then, by [25, Cor. 1.4.3], the sequence $(u_n x)_{n \in \mathbb{N}_0}$ is uniformly distributed modulo 1 for a.e. $x \in \mathbb{R}$. In fact, it is a rather direct consequence that, for any $k \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ and any real number $L > 0$, the arithmetic progression sequence $(u_{km+\ell} x)_{m \in \mathbb{N}}$ is uniformly distributed modulo L for a.e. $x \in \mathbb{R}$. This is the *total Bohr ergodicity* of the sequence $(u_n)_{n \in \mathbb{N}_0}$ as introduced in [18, Def. 2.1]. Clearly, $u_n = \alpha^n$ with $|\alpha| > 1$ defines such a sequence, while no bounded sequence can have this property. \diamond

As soon as we leave the realm of periodic functions that are locally Riemann-integrable, the desired averaging statements will need some finer properties of our sequences $(\alpha^n x)_{n \in \mathbb{N}}$, where we assume $|\alpha| > 1$ as before. In particular, we will need details on the uniform distribution (or the deviation from it) and some information on the approximation or non-approximation of numbers in a given set by the sequence elements. For the first issue, we need the discrepancy structure of the sequence, and some Diophantine approximation properties for the latter.

Recall that the *discrepancy* of a sequence $(x_n)_{n \in \mathbb{N}}$ is quantified in terms of the first N elements of the sequence (taken modulo 1), namely by the number

$$\mathcal{D}_N := \sup_{0 \leq a < b \leq 1} \left| \frac{1}{N} \text{card}([a, b) \cap \{\langle x_1 \rangle, \dots, \langle x_N \rangle\}) - (b - a) \right|,$$

together with its asymptotic properties as $N \rightarrow \infty$.

In this terminology, the sequence $(x_n)_{n \in \mathbb{N}}$ is uniformly distributed modulo 1 if and only if $\lim_{N \rightarrow \infty} \mathcal{D}_N = 0$; see [11, Thm. 1.12]. The discrepancy of sequences is a huge research area in itself; see [17] and references therein for a comprehensive exposition. We will need the following result, where we refer to [20, Thm. 5.13] and [23] for proofs.

Fact 2.5. *Let $\alpha \in \mathbb{R}$ with $|\alpha| > 1$ be given. Then, for any fixed $\varepsilon > 0$, the discrepancy of the sequence $(\alpha^n x)_{n \in \mathbb{N}}$, for a.e. $x \in \mathbb{R}$, asymptotically is*

$$\mathcal{D}_N = \mathcal{O} \left(\frac{(\log(N))^{\frac{3}{2} + \varepsilon}}{\sqrt{N}} \right)$$

as $N \rightarrow \infty$. \square

Next, we need a *Diophantine approximation* property. If $\emptyset \neq Y \subset \mathbb{R}$ is a uniformly discrete point set, compare [3, Sec. 2.1], we can define

$$\text{dist}(x, Y) := \min_{y \in Y} |x - y|$$

as the distance of $x \in \mathbb{R}$ from Y . Now, one can state the following metric ‘non-approximation’ result, which is a versatile generalisation of the classic situation with $Y = \mathbb{Z}$.

Lemma 2.6. *Let $\alpha \in \mathbb{R}$ with $|\alpha| > 1$ be given, and let $Y \subset \mathbb{R}$ be a non-empty, uniformly discrete point set. Further, fix some $\varepsilon > 0$. Then, for a.e. $x \in \mathbb{R}$, the inequality*

$$\text{dist}(\alpha^{n-1}x, Y) \geq \frac{1}{n^{1+\varepsilon}}$$

holds for almost all $n \in \mathbb{N}$, by which we mean that it holds for all natural numbers except at most finitely many.

Proof. The statement is trivial when Y is a finite point set, so let us assume that Y is unbounded. In this case, one still has

$$\delta := \inf \{|x - y| : x, y \in Y, x \neq y\} > 0,$$

due to the assumed uniform discreteness of Y . Consequently, the number of points of Y in an arbitrary interval $[a, b]$ with $a \leq b$ satisfies

$$(2.1) \quad \text{card}(Y \cap [a, b]) \leq 1 + \left\lceil \frac{b - a}{\delta} \right\rceil,$$

where $[\cdot]$ is the Gauß bracket.

Let $m \in \mathbb{Z}$ be arbitrary, but fixed, and consider $I_m = [m, m+1]$. With $\mathbb{R} = \bigcup_{m \in \mathbb{Z}} I_m$, it suffices to show that our claim fails at most for a null set within the interval I_m , as the countable union of null sets is still a null set.

Choose $\varepsilon > 0$ and, for $n \in \mathbb{N}$, consider the set

$$A_n^{(m)} := \left\{ x \in I_m : \text{dist}(\alpha^{n-1}x, Y) < \frac{1}{n^{1+\varepsilon}} \right\}.$$

It is clearly measurable, and its measure, since $|\alpha| > 1$, can be estimated as

$$\begin{aligned} \lambda(A_n^{(m)}) &= \frac{1}{|\alpha|^{n-1}} \lambda \left\{ z \in \alpha^{n-1}I_m : \text{dist}(z, Y) < \frac{1}{n^{1+\varepsilon}} \right\} \\ &\leq \frac{1}{|\alpha|^{n-1}} \frac{2}{n^{1+\varepsilon}} \left(1 + \left\lceil \frac{|\alpha|^{n-1}}{\delta} \right\rceil \right) = \mathcal{O}\left(\frac{1}{n^{1+\varepsilon}}\right), \end{aligned}$$

where the second step is a consequence of Eq. (2.1). We thus know that there is a $C > 0$ such that $\lambda(A_n^{(m)}) \leq C/n^{1+\varepsilon}$ for all $n \in \mathbb{N}$.

Now, we have

$$0 \leq \sum_{n \geq 1} \lambda(A_n^{(m)}) \leq C \sum_{n \geq 1} \frac{1}{n^{1+\varepsilon}},$$

where the second sum is convergent, and thus also the first. Then, Cantelli's lemma² tells us that

$$E_\infty^{(m)} := \{x \in I_m : x \in A_n^{(m)} \text{ for infinitely many } n \in \mathbb{N}\}$$

is indeed a null set, which is what we needed to show. \square

Remark 2.7. Though immaterial for the proof, it is often useful in an application to also remove all $x \in \mathbb{R}$ with $Y \cap \{\alpha^{n-1}x : n \in \mathbb{N}\} \neq \emptyset$, which constitutes a null set because it is clearly countable or even finite. \diamond

The lower bound in Lemma 2.6 can be replaced by the values of a more general, non-negative arithmetic function, $\psi(n)$ say, provided one also has the summability condition $\sum_{n \in \mathbb{N}} \psi(n) < \infty$. When this sum diverges, the situation changes. Indeed, for instance if $\alpha = 2$ and $Y = \mathbb{Z}$, there is then a set $X \subset \mathbb{R}$ of full measure such that, for $x \in X$, the distance of $2^{n-1}x$ from the nearest integer is smaller than $\frac{1}{n}$ for infinitely many $n \in \mathbb{N}$; see [4] for a more general result in this direction. Moreover, one cannot do better than using some $\varepsilon > 0$ in Lemma 2.6, in line with the divergence of the harmonic series.

3. AVERAGING PERIODIC FUNCTIONS

Let us first state a result that emerges from an application of Weyl's criterion to the special type of sequences we are interested in.

Fact 3.1. *Let $f: \mathbb{R} \rightarrow \mathbb{C}$ be a continuous or, more generally, a locally Riemann-integrable function that is L -periodic, so $f(x+L) = f(x)$ holds for some fixed $L > 0$ and all $x \in \mathbb{R}$. If α is a real number with $|\alpha| > 1$, one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) = \frac{1}{L} \int_0^L f(y) dy$$

for a.e. $x \in \mathbb{R}$.

Proof. Since any L -periodic continuous function is also locally Riemann-integrable, it suffices to consider the latter class. Define a new function g by $g(x) := f(Lx)$, which clearly is 1-periodic and locally Riemann-integrable. Now, we have

$$\frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) = \frac{1}{N} \sum_{n=0}^{N-1} g(\alpha^n \frac{x}{L}),$$

where $(\alpha^n \frac{x}{L})_{n \in \mathbb{N}}$ is uniformly distributed modulo 1 for a.e. $\frac{x}{L} \in \mathbb{R}$, and hence also for a.e. $x \in \mathbb{R}$, by Fact 2.3. Consequently, Weyl's criterion from Lemma 2.1 tells us that

$$\frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) \xrightarrow{N \rightarrow \infty} \int_0^1 g(z) dz = \frac{1}{L} \int_0^L f(y) dy$$

holds for all such cases, which means for a.e. $x \in \mathbb{R}$ as claimed. \square

²This being the 'easy half' of the Borel–Cantelli lemma, which goes back to Cantelli, we follow [11, App. C] in our terminology, and also refer to this reference for a proof.

Note that one can rewrite Fact 3.1 with the *mean* of f , because

$$M(f) := \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(y) \, dy = \lim_{T \rightarrow \infty} \frac{1}{T} \int_a^{a+T} f(y) \, dy = \frac{1}{L} \int_0^L f(y) \, dy$$

holds for every L -periodic function that is locally Riemann-integrable, where the limit clearly is uniform in $a \in \mathbb{R}$.

Example 3.2. Fix $k \in \mathbb{R}$ and consider the trigonometric monomial defined by $\psi_k(x) = e^{2\pi i k x}$. Unless $k = 0$, in which case $\psi_0 \equiv 1$, the function ψ_k has period $\frac{1}{|k|} > 0$. For $\alpha \in \mathbb{R}$ with $|\alpha| > 1$, Fact 3.1 implies that

$$\frac{1}{N} \sum_{n=0}^{N-1} \psi_k(\alpha^n x) \xrightarrow{N \rightarrow \infty} M(\psi_k) = \begin{cases} 1, & k = 0, \\ 0, & \text{otherwise,} \end{cases}$$

holds for a.e. $x \in \mathbb{R}$.

More generally, if $(u_n)_{n \in \mathbb{N}_0}$ with $\inf_{n \neq m} |u_n - u_m| > 0$ is the type of sequence from Remark 2.4, the above convergence statement also holds with $\alpha^n x$ replaced by $u_n x$. \diamond

It is clear from the proof of Fact 3.1 that, for periodic functions, it suffices to consider the case $L = 1$ without loss of generality, as we do from now on. Our next step shows that, for $\alpha \in \mathbb{Z}$, one can go beyond the class of 1-periodic functions that are locally Riemann-integrable.

Lemma 3.3. *Consider a function $f \in L^1_{\text{loc}}(\mathbb{R})$ that is 1-periodic. Fix $q \in \mathbb{Z}$ with $|q| \geq 2$. Then, for a.e. $x \in \mathbb{R}$, one has*

$$\frac{1}{N} \sum_{n=0}^{N-1} f(q^n x) \xrightarrow{N \rightarrow \infty} \int_0^1 f(y) \, dy = M(f).$$

Proof. Since $q \in \mathbb{Z}$, we may view the average as a Birkhoff sum for the dynamical system on $[0, 1]$ defined by the mapping $x \mapsto qx \bmod 1$. It is well known that Lebesgue measure is invariant and ergodic for this system, compare [12] and references therein, wherefore we may employ Birkhoff's ergodic theorem [38] to f , which is Lebesgue-integrable on $[0, 1]$ by assumption, and our claim follows. \square

Note that the exceptional set, for which the limit differs or does not exist, may depend on f when the latter fails to be continuous. In fact, there clearly is no uniformly distributed sequence that will work for *all* 1-periodic $f \in L^1_{\text{loc}}(\mathbb{R})$. Still, the result of Lemma 3.3 suggests that something more general than Fact 3.1 might also be true when our multiplier α fails to be an integer. However, we cannot apply the ‘trick’ with Birkhoff's ergodic theorem when $\alpha \notin \mathbb{Z}$. This is due to the fact that the sequence $(\langle \alpha^n x \rangle)_{n \in \mathbb{N}_0}$, which is uniformly distributed on $[0, 1)$ for a.e. $x \in \mathbb{R}$ by Fact 2.3, does no longer agree with the orbit of x under the mapping T defined by $x \mapsto \alpha x \bmod 1$. The latter, for a.e. $x \in \mathbb{R}$, follows the distribution of the (ergodic) Rényi–Parry measure [32, 30] for α , which is of the form $h_\alpha \lambda$ with h_α being Lebesgue-integrable on $[0, 1)$. When $\alpha \notin \mathbb{Z}$, the measures λ and $h_\alpha \lambda$ are still equivalent as measures, but different; see [12] and references therein for more.

Example 3.4. To illustrate the difference, consider $\alpha = \tau = \frac{1}{2}(1 + \sqrt{5})$, which is one of the simplest examples in this context. When f is 1-periodic and locally Riemann-integrable, we get

$$\frac{1}{N} \sum_{n=0}^{N-1} f(\tau^n x) \xrightarrow{N \rightarrow \infty} \int_0^1 f(x) dx = M(f)$$

for a.e. $x \in \mathbb{R}$ by Weyl's criterion (Lemma 2.1).

In comparison, let T be defined by $x \mapsto \tau x \bmod 1$ on $[0, 1)$. Then, for a.e. $x \in [0, 1)$, the orbits $(T^n x)_{n \in \mathbb{N}_0}$ follow the distribution given by the piecewise constant function [32, Ex. 4]

$$h_\tau(x) = \begin{cases} \frac{5+3\sqrt{5}}{10}, & 0 \leq x < \frac{1}{\tau}, \\ \frac{5+\sqrt{5}}{10}, & \frac{1}{\tau} \leq x < 1. \end{cases}$$

Since T is ergodic for the measure $h_\tau \lambda$, Birkhoff's theorem tells us that, for any Lebesgue-integrable function f on $[0, 1)$, one has

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \int_0^1 f(x) h_\tau(x) dx$$

for a.e. $x \in [0, 1)$, and this limit will generally differ from $M(f)$.

Moreover, since the sequences $(\langle \tau^n x \rangle)_{n \geq 0}$ and $(T^n x)_{n \geq 0}$ are not easily relatable, one cannot infer the convergence of averages along the exponential sequence from those along the orbits under T . \diamond

Let us now extend Fact 3.1 beyond Riemann-integrable functions by stating one version of Sobol's theorem [34, Thm. 1].

Theorem 3.5. *Let $\alpha \in \mathbb{R}$ with $|\alpha| > 1$ be fixed, and consider a 1-periodic function $f \in L^1_{\text{loc}}(\mathbb{R})$ that fails to be locally Riemann-integrable. Assume that there is a finite set $F \subset [0, 1]$ such that f , for every $\delta > 0$, is Riemann-integrable on the complement of $F + (-\delta, \delta)$ in $[0, 1]$. Assume further that, for every $z \in F$, there is a $\delta_z > 0$ such that f is differentiable on the punctured interval $(z - \delta_z, z + \delta_z) \setminus \{z\}$ and that, for any $s > 0$,*

$$V_N(z, s) := \int_{z-\delta_z}^{z-\frac{1}{N^s}} |f'(x)| dx + \int_{z+\frac{1}{N^s}}^{z+\delta_z} |f'(x)| dx = \mathcal{O}(N^{\frac{s}{2}-\eta})$$

holds for some $\eta = \eta(z) > 0$ as $N \rightarrow \infty$.

Then, for a.e. $x \in \mathbb{R}$, one has

$$\frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) \xrightarrow{N \rightarrow \infty} \int_0^1 f(y) dy = M(f).$$

Sketch of proof. Since F is finite, we may choose $0 < \delta \leq \min_{z \in F} \delta_z$ small enough such that the open sets $(z - \delta, z + \delta)$ with $z \in F$ are disjoint. By writing f as a sum of a locally Riemann-integrable function (such as the restriction $f^{(\delta)}$ of f to the complement of $\mathbb{Z} + F + (-\delta, \delta)$) and $r = \text{card}(F)$ ‘problematic’ terms, the latter supported on $(z - \delta, z + \delta)$ with $z \in F$, it is clear that our claim follows if we can deal with one of these problematic terms. So, select

one $z \in F$. Without loss of generality, we may assume that $(z - \delta, z + \delta) \subset [0, 1]$, as we can otherwise shift the unit interval because f and $f^{(\delta)}$ are 1-periodic.

One can now repeat the original proof from [34], or the more extensive version in [21, Sec. 2]. Here, the validity of the convergence claim emerges from the observation that, for a.e. $x \in \mathbb{R}$, the number $\langle \alpha^{n-1} x \rangle$ does not come closer to z than $1/n^{1+\varepsilon}$, for any fixed $\varepsilon > 0$ and then all $n \in \mathbb{N}$ except at most finitely many. This follows from Lemma 2.6 with $Y = z + \mathbb{Z}$. Now,

$$V_N(z, 1 + \varepsilon) = o\left(N^{\frac{1+\varepsilon}{2}-\eta}\right)$$

for some $\eta > 0$ by assumption. Since η does not depend on ε , we are still free to choose $\varepsilon > 0$ small enough so that $\vartheta := \eta - \frac{\varepsilon}{2} > 0$.

Now, the potentially large contribution to our averaging sum from sequence elements close to z are properly ‘counterbalanced’ by the discrepancy of $(\alpha^n x)_{n \in \mathbb{N}}$, where we invoke Fact 2.5 with the ε just chosen. One obtains

$$(3.1) \quad \mathcal{D}_N \cdot V_N(z, 1 + \varepsilon) = o\left(\frac{(\log(N))^{\frac{3}{2}+\varepsilon}}{N^\vartheta}\right) = o(1),$$

which is a sufficient criterion for the claimed convergence because

$$\frac{1}{N} \sum_{n=0}^{N-1} f^{(\delta)}(\alpha^n x) \xrightarrow{N \rightarrow \infty} M(f^{(\delta)})$$

holds for a.e. $x \in \mathbb{R}$, while the Birkhoff average of $f - f^{(\delta)}$ is controlled by Eq. (3.1) and tends to 0 as $\delta \searrow 0$. \square

Remark 3.6. The assumption that F in Theorem 3.5 is a finite set implies $\delta := \min_{z \in F} \delta_z > 0$ as well as $\min_{z \in F} \eta(z) > 0$. Later, we will replace this setting by a suitable compactness assumption to extend the result of this theorem to almost periodic functions. \diamond

Remark 3.7. The differentiability assumption for f near the ‘bad’ points is convenient, but not necessary. It can be replaced by the requirement that the total variation of f on sets of the form $(z - \delta, z - N^{-s}] \cup [z + N^{-s}, z + \delta)$ behaves as stated for $V_N(z, s)$; compare [34, 21]. \diamond

As mentioned earlier, results of this type are also of interest for the numerical calculation of integrals, for instance with methods of (quasi-) Monte Carlo type. In our context, an important question is how to extend Riesz–Raikov sums and Birkhoff averages to functions that fail to be periodic, but possess some repetitivity structure instead.

4. AVERAGING ALMOST PERIODIC FUNCTIONS

At this point, we need to recall some basic definitions and results from the theory of almost periodic functions in the sense of Bohr [9], where we refer to [3, Sec. 8.2] for a short summary, to [24, Sec. VI.5] or [14] for comprehensive expositions, and to [28, Sec. 41] for a more general and abstract setting (including non-Abelian groups).

Recall that $f \in C(\mathbb{R})$ is called *almost periodic in the sense of Bohr* if, for any $\varepsilon > 0$, the set of ε -almost periods

$$\mathcal{P}_\varepsilon := \{t \in \mathbb{R} : \|f - T_t f\|_\infty < \varepsilon\}$$

is relatively dense in \mathbb{R} . Here, $(T_t f)(x) := f(x-t)$ defines the t -translate of f . Any continuous periodic function is almost periodic in this sense, as is any trigonometric polynomial. Any Bohr-almost periodic function is bounded and uniformly continuous. In fact, the $\|\cdot\|_\infty$ -closure of the (complex) algebra of trigonometric polynomials is precisely the space of *all* Bohr-almost periodic functions [9].

For comparison, $f \in C(\mathbb{R})$ is called *almost periodic in the sense of Bochner* (for $\|\cdot\|_\infty$, to be precise) if the translation orbit $\{T_t f : t \in \mathbb{R}\}$ is precompact in the $\|\cdot\|_\infty$ -topology. The fundamental relation among these notions can be summarised as follows; see [3, Prop. 8.2] as well as [24, 14].

Fact 4.1. *For $f \in C(\mathbb{R})$, the following properties are equivalent.*

- (1) *f is Bohr-almost periodic, i.e., \mathcal{P}_ε is relatively dense for any $\varepsilon > 0$;*
- (2) *f is Bochner-almost periodic for $\|\cdot\|_\infty$, i.e., the orbit $\{T_t f : t \in \mathbb{R}\}$ is precompact in the $\|\cdot\|_\infty$ -topology;*
- (3) *f is the limit of a sequence of trigonometric polynomials, with uniform convergence of the sequence on \mathbb{R} .* □

In view of these relations, we follow [6] and speak of *uniformly almost periodic functions* from now on when we refer to this class. If misunderstandings are unlikely, we will drop the attribute ‘uniformly’. Let us elaborate a little on part (3) of Fact 4.1. If f is almost periodic, its mean

$$(4.1) \quad M(f) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{a-T}^{a+T} f(x) \, dx$$

exists for any $a \in \mathbb{R}$, is independent of a , and the convergence is uniform in a ; compare [36] for a more detailed discussion of this concept. When we need to emphasise the role of a for more general types of functions (say without uniformity of the limit in a), we will write $M(f; a)$.

The *Fourier–Bohr coefficient* of an almost periodic function f at $k \in \mathbb{R}$ is given by

$$a(k) = M(e^{-2\pi i k(\cdot)} f).$$

It exists for any $k \in \mathbb{R}$, and differs from 0 for at most countably many values of k . Any $k \in \mathbb{R}$ with $a(k) \neq 0$ is called a *frequency* of f . If $\{k_\ell\}$ is the set of frequencies of f , there is a sequence of trigonometric polynomials of the form

$$(4.2) \quad P^{(m)}(x) = \sum_{\ell=1}^{n_m} r_\ell^{(m)} a(k_\ell) e^{2\pi i k_\ell x}$$

that converge uniformly to f on \mathbb{R} as $m \rightarrow \infty$. Here, the numbers $r_\ell^{(m)}$, which are known as convergence enforcing numbers, depend on m and k_ℓ , but not on $a(k_\ell)$, and can be chosen as rational numbers [14, Thm. I.1.24].

To approach averages of almost periodic functions, it is thus more than natural to begin with the averages of trigonometric polynomials. We formulate the next result for more general sequences than the exponential ones from above.

Proposition 4.2. *Let P_m be a (complex) trigonometric polynomial of the form*

$$P_m(x) = a_0 + \sum_{\ell=1}^m a_\ell e^{2\pi i k_\ell x},$$

with coefficients $a_\ell \in \mathbb{C}$ and distinct non-zero frequencies k_1, \dots, k_m . Further, let $(u_n)_{n \in \mathbb{N}_0}$ be a sequence of real numbers such that $\inf_{n \neq m} |u_n - u_m| > 0$. Then, for a.e. $x \in \mathbb{R}$, one has

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} P_m(u_n x) = M(P_m) = a_0.$$

In particular, this holds for $u_n = \alpha^n$ with $\alpha \in \mathbb{R}$ and $|\alpha| > 1$.

Proof. The claim is obvious for $m = 0$, where the polynomial is constant. The case $m = 1$ with $a_0 = 0$, where P_m is a monomial, is Example 3.2 from above. So, for a general P_m , the claim is true for each summand individually, with an exceptional set $E(k_\ell)$ of measure 0 for $\ell \geq 1$. Since $\bigcup_{\ell=1}^m E(k_\ell)$ is still a null set, the statement on the limit is clear, while its value follows from a simple calculation with the mean; compare Example 3.2. \square

Before we proceed, let us recall the following useful property of the mean.

Lemma 4.3. *Let $(g_n)_{n \in \mathbb{N}}$ be a sequence of complex-valued, but not necessarily continuous, functions on \mathbb{R} that converge uniformly to a function f . Assume further that the mean $M(g_n)$ exists for all $n \in \mathbb{N}$. Then, also $M(f)$ exists, and $\lim_{n \rightarrow \infty} M(g_n) = M(f)$. In particular, one has*

$$M(f) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{a-T}^{a+T} f(x) dx$$

for any fixed $a \in \mathbb{R}$. When the convergence of the means $M(g_n) = M(g_n; a)$ is uniform in a , then so is the convergence of $M(f)$.

Proof. The assumed uniform convergence also means that $(g_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in the $\|\cdot\|_\infty$ -topology. Fix $\varepsilon > 0$ and choose $n_0 = n_0(\varepsilon)$ such that $\|g_n - f\|_\infty < \varepsilon$ as well as $\|g_n - g_m\|_\infty < \varepsilon$ holds for all $n, m \geq n_0$. Then, for any $T > 0$, one has

$$(4.3) \quad \frac{1}{2T} \left| \int_{-T}^T (g_n(x) - g_m(x)) dx \right| \leq \|g_n - g_m\|_\infty < \varepsilon$$

for all $n, m \geq n_0$, which implies

$$\begin{aligned} |M(g_n) - M(g_m)| &\leq \|g_n - g_m\|_\infty \\ &+ \left| M(g_n) - \frac{1}{2T} \int_{-T}^T g_n(x) dx \right| + \left| M(g_m) - \frac{1}{2T} \int_{-T}^T g_m(x) dx \right|. \end{aligned}$$

Consequently, $|M(g_n) - M(g_m)| < 3\varepsilon$ for all sufficiently large T due to our assumption on the existence of the means $M(g_n)$. Note that, although T may depend on m and n , the

above 3ε -estimate still works as a consequence of Eq. (4.3). The sequence $(M(g_n))_{n \in \mathbb{N}}$ is thus Cauchy, hence convergent, with limit \mathfrak{M} , say.

Now, choose $n \geq n_0$ large enough such that also $|M(g_n) - \mathfrak{M}| < \varepsilon$ holds, fix an arbitrary $a \in \mathbb{R}$, and consider

$$\begin{aligned} \left| \frac{1}{2T} \int_{a-T}^{a+T} f(x) dx - \mathfrak{M} \right| &\leq \|f - g_n\|_\infty + |M(g_n) - \mathfrak{M}| \\ &\quad + \left| \frac{1}{2T} \int_{a-T}^{a+T} g_n(x) dx - M(g_n) \right| < 3\varepsilon, \end{aligned}$$

where the last step holds for all sufficiently large T by assumption. This derivation implies

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{a-T}^{a+T} f(x) dx = \mathfrak{M} = \lim_{n \rightarrow \infty} M(g_n),$$

which is independent of $a \in \mathbb{R}$, and $M(f) = \mathfrak{M}$ is the claimed mean of f .

When, in addition, the means of the functions g_n exist uniformly in a , our 3ε -argument also implies that the convergence of $M(f; a)$ is uniform in $a \in \mathbb{R}$ as claimed. \square

This enables us to formulate the following result.

Theorem 4.4. *Let $\alpha \in \mathbb{R}$ with $|\alpha| > 1$ be given, and let f be a Bohr-almost periodic function on \mathbb{R} . Then, for a.e. $x \in \mathbb{R}$, one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) = M(f).$$

Proof. Let $(g_n)_{n \in \mathbb{N}}$ be a sequence of trigonometric polynomials that converge uniformly to f . As is well known, compare [14], and is a rather direct consequence of Eq. (4.2), the sequence can be chosen such that the frequency sets $\{k_j : 1 \leq j \leq m_n\}$ of the g_n are nested. By Proposition 4.2, we know that, for every $n \in \mathbb{N}$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\ell=0}^{N-1} g_n(\alpha^\ell x) = M(g_n)$$

holds for a.e. $x \in \mathbb{R}$, where we denote the excluded null set by E_n . By construction, we have $E_n \subseteq E_{n+1}$, and $E := \bigcup_{n \in \mathbb{N}} E_n$ is still a null set.

Define the Birkhoff average of φ at x as $S_N(\varphi, x) = \frac{1}{N} \sum_{n=0}^{N-1} \varphi(\alpha^n x)$, and fix some $\varepsilon > 0$. Choose $n_0 = n_0(\varepsilon)$ such that $\|f - g_n\|_\infty < \varepsilon$ for all $n \geq n_0$, which is possible under our assumptions. Now, for any fixed $x \in \mathbb{R} \setminus E$, we can estimate

$$|S_N(f, x) - M(f)| \leq |S_N(f - g_n, x)| + |S_N(g_n, x) - M(g_n)| + |M(g_n) - M(f)|$$

where, independently of N ,

$$|S_N(f - g_n, x)| \leq S_N(|f - g_n|, x) \leq \|f - g_n\|_\infty < \varepsilon$$

for any $n \geq n_0$. The third term on the right-hand side of the previous estimate is smaller than ε for sufficiently large n as a consequence of Lemma 4.3, while the middle term, under

our assumptions, is bounded by ε for sufficiently large N , which we are still free to choose. This 3ε -argument thus establishes the claim. \square

Let us mention in passing that Theorem 4.4 still holds if α^n , as before, is replaced by the numbers u_n of a sequence as described in Remark 2.4.

At this point, to go any further, we need to extend the class of functions we consider. This is motivated by the fact that uniform almost periodicity is often too restrictive. In particular, in various examples from dynamical systems theory, one encounters averages over functions that fail to be bounded, and hence cannot be uniformly almost periodic. Being unbounded, such functions cannot be locally Riemann-integrable either, though they might still admit improper Riemann integrals or be locally Lebesgue-integrable.

It would be natural to investigate the question in the setting of weakly almost periodic functions, as introduced in [36], which seems possible as well. However, the above remarks indicate that one needs results also for functions that violate continuity. This suggests to use the wider class of almost periodic functions in the sense of Stepanov³ [35], which relate to locally Lebesgue-integrable functions like uniform almost periodic functions do to continuous functions. The new norm on $L^1_{\text{loc}}(\mathbb{R})$ is given by

$$\|f\|_S = \sup_{x \in \mathbb{R}} \frac{1}{L} \int_x^{x+L} |f(y)| dy,$$

where $L > 0$ is an arbitrary, but fixed number. Since these norms are equivalent for different values of L , it is most convenient to choose $L = 1$, as we do from now on. Now, a locally Lebesgue-integrable function f is called *almost periodic in the sense of Stepanov*, or S-almost periodic for short, if, for any $\varepsilon > 0$, the set $\mathcal{P}_\varepsilon^S$ of ε -almost periods of f for $\|\cdot\|_S$ is relatively dense. The analogue of Fact 4.1 then reads as follows (we omit a proof because it works the same way as in the previous case; compare [14]).

Fact 4.5. *For $f \in L^1_{\text{loc}}(\mathbb{R})$, the following properties are equivalent.*

- (1) *f is S-almost periodic, i.e., $\mathcal{P}_\varepsilon^S$ is relatively dense for any $\varepsilon > 0$;*
- (2) *f is Bochner-almost periodic for $\|\cdot\|_S$, i.e., the orbit $\{T_t f : t \in \mathbb{R}\}$ is precompact in the $\|\cdot\|_S$ -topology;*
- (3) *f is the $\|\cdot\|_S$ -limit of a sequence of trigonometric polynomials.* \square

Let us note in passing that every locally integrable function f on \mathbb{R} may be viewed as a translation bounded measure (where f is the Radon–Nikodym density relative to λ). In doing so, the Stepanov norm is induced by the $\|\cdot\|_{[0,1]}$ -norm for measures as discussed in [36]. This implies that a function $f \in L^1_{\text{loc}}(\mathbb{R})$ is S-almost periodic if and only if the measure $f\lambda$ is norm-almost periodic in the sense of [5, 36].

Every uniformly almost periodic function is S-almost periodic, which also means (via part (3) of Fact 4.5) that any S-almost periodic function can be $\|\cdot\|_S$ -approximated by uniformly

³The widely used modern version of the name is V.V. Stepanov, while the author used W. Stepanoff in his original articles.

almost periodic functions. In other words, the class of all S-almost periodic functions can equivalently be described as the $\|\cdot\|_S$ -closure of the (complex) algebra of trigonometric polynomials or as that of the class of uniformly almost periodic functions. Moreover, the space of S-almost periodic functions is complete in the $\|\cdot\|_S$ -norm, and $\|f\|_S = 0$ means $f = 0$ in the Lebesgue sense, so $f(x) = 0$ for a.e. $x \in \mathbb{R}$; see [7] for details.

Remark 4.6. If f is S-almost periodic, its mean exists. In fact, observe that, for all S-almost periodic functions f, g and for any $a \in \mathbb{R}$, one has

$$\frac{1}{2T} \left| \int_{a-T}^{a+T} (f(x) - g(x)) \, dx \right| \leq \frac{1 + [2T]}{2T} \|f - g\|_S.$$

Now, it is immediate that the statement of Lemma 4.3 still holds if uniform convergence is replaced by $\|\cdot\|_S$ -convergence. This then gives the desired existence of means because, by Fact 4.5(3), we can $\|\cdot\|_S$ -approximate any S-almost periodic function with trigonometric polynomials for which the mean clearly exists. \diamond

As an aside, we mention the following interesting connection.

Lemma 4.7. *Let f be an S-almost periodic function, and let $\delta > 0$ be arbitrary, but fixed. Then, the function f_δ defined by*

$$f_\delta(x) = \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} f(y) \, dy$$

is continuous and uniformly almost periodic. Moreover, $\lim_{\delta \searrow 0} f_\delta = f$ in the $\|\cdot\|_S$ -topology.

Proof. Assume $\delta \leq \frac{1}{2}$ (the argument for $\delta > \frac{1}{2}$ is analogous), and let t be a $(2\delta\varepsilon)$ -almost period of f for $\|\cdot\|_S$. Now,

$$\begin{aligned} |f_\delta(t+x) - f_\delta(x)| &= \frac{1}{2\delta} \left| \int_{x-\delta}^{x+\delta} f(t+y) - f(y) \, dy \right| \\ &\leq \frac{1}{2\delta} \int_{x-\delta}^{x+\delta+1} |f(t+y) - f(y)| \, dy \leq \frac{\|f - T_t f\|_S}{2\delta} < \varepsilon \end{aligned}$$

which implies that t is an ε -almost period of f_δ for $\|\cdot\|_\infty$. Via part (1) of Fact 4.5, we conclude that f_δ satisfies part (1) of Fact 4.1, and thus is uniformly almost periodic. As such, f_δ is also uniformly continuous.

For the second claim, we refer to the original proof in [7], which uses an approximation argument that is based on the effect that a ‘convolution mollifier’ has on a locally Lebesgue-integrable function. \square

The main extension of Theorem 3.5 can be stated as follows.

Theorem 4.8. *Let $\alpha \in \mathbb{R}$ with $|\alpha| > 1$ be fixed, and let $f \in L^1_{\text{loc}}(\mathbb{R})$ be an S-almost periodic function. Assume now that there is a uniformly discrete set $Y \subset \mathbb{R}$ such that f , for every $\delta > 0$, is locally Riemann-integrable on the complement of $Y + (-\delta, \delta)$. Assume further that*

there is a $\delta' > 0$ such that, for any $z \in Y$, f is differentiable on the punctured interval $(z - \delta', z + \delta') \setminus \{z\}$ and that, for any $s > 0$ and with $V_N(z, s)$ as defined in Theorem 3.5,

$$\sup_{z \in Y} V_N(z, s) = o(N^{\frac{s}{2} - \eta})$$

holds for some $\eta > 0$ as $N \rightarrow \infty$.

Then, for a.e. $x \in \mathbb{R}$, one has

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) = M(f),$$

where the mean exists because f is S-almost periodic.

Sketch of proof. Without loss of generality, we may assume that δ' is small enough so that the open intervals $(z - \delta', z + \delta')$ with $z \in Y$ are disjoint. Now, Lemma 2.6 guarantees that the sequence $(\alpha^{n-1}x)_{n \in \mathbb{N}}$, for a.e. $x \in \mathbb{R}$, does not come closer to Y than $1/n^{1+\varepsilon}$, for any fixed $\varepsilon > 0$ and then for all $n \in \mathbb{N}$ except at most finitely many.

For any $z \in Y$, we have $V_N(z, 1 + \varepsilon) = o(N^{\frac{1+\varepsilon}{2} - \eta})$ for some fixed $\eta > 0$ by assumption, where we may once again assume that $\varepsilon > 0$ is chosen such that $\vartheta = \eta - \frac{\varepsilon}{2} > 0$. With the estimate of Eq. (3.1) in the proof of Theorem 3.5, we again obtain $\mathcal{D}_N \cdot V_N(z, 1 + \varepsilon) = o(1)$ as $N \rightarrow \infty$, which establishes a sufficient criterion for the claimed convergence.

Indeed, let $0 < \delta < \delta'$ be arbitrary, and let 1_δ denote the characteristic function of the set $\mathbb{R} \setminus (Y + (-\delta, \delta))$. Obviously, for any such δ , the function $f^{(\delta)} := f \cdot 1_\delta$ is both S-almost periodic and locally Riemann-integrable on \mathbb{R} . For a.e. $x \in \mathbb{R}$, we thus get

$$\frac{1}{N} \sum_{n=0}^{N-1} f^{(\delta)}(\alpha^n x) \xrightarrow{N \rightarrow \infty} M(f^{(\delta)}) \xrightarrow{\delta \searrow 0} M(f)$$

by a combination of our previous arguments. Since the average of $f - f^{(\delta)}$ along the exponential sequence is controlled by the above mentioned estimate from Eq. (3.1), our claim follows. \square

Note that our assumption on η achieves the analogue of the comment made in Remark 3.6. Note also that Remark 3.7 has an obvious extension to this more general situation. Indeed, one can once again replace the differentiability condition by the corresponding behaviour of the total variation in the vicinity of the ‘bad’ points.

5. FURTHER DIRECTIONS AND EXTENSIONS

Our exposition so far used complex-valued almost periodic functions over \mathbb{R} , mainly for ease of presentation. More generally, one is interested in vector-valued functions, or in function with values in an arbitrary Banach space \mathbb{X} , with norm $|\cdot|$ say. So, let $f: \mathbb{R} \rightarrow \mathbb{X}$ be such a function, and define $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$. Then, the ε -almost periods of f are again defined as

$$\mathcal{P}_\varepsilon := \{t \in \mathbb{R} : \|f - T_t f\|_\infty < \varepsilon\},$$

with $(T_t f)(x) = f(x - t)$ as before.

Likewise, one can define trigonometric polynomials (or functions), by which one now means any function $T: \mathbb{R} \rightarrow \mathbb{X}$ of the form

$$(5.1) \quad Q_m(x) = a_0 + \sum_{\ell=1}^m e^{2\pi i k_\ell x} a_\ell$$

for some $m \geq 0$, where $\{k_1, \dots, k_\ell\}$ are distinct, non-zero real numbers and where the a_ℓ are now elements of \mathbb{X} . When $m = 0$, the sum is meant to be empty and Q_m is constant. The analogue of Fact 4.1 can now be stated as follows; see [14, Ch. VI] for details.

Fact 5.1. *Let $(\mathbb{X}, |\cdot|)$ be a Banach space. Then, for a continuous function $f: \mathbb{R} \rightarrow \mathbb{X}$, the following properties are equivalent.*

- (1) *f is Bohr-almost periodic, i.e., \mathcal{P}_ε is relatively dense for any $\varepsilon > 0$;*
- (2) *f is Bochner-almost periodic for $\|\cdot\|_\infty$, i.e., the orbit $\{T_t f : t \in \mathbb{R}\}$ is precompact in the $\|\cdot\|_\infty$ -topology;*
- (3) *f is the limit of a sequence of trigonometric polynomials, with uniform convergence of the sequence on \mathbb{R} .* □

There is no surprise up to this point, and we have gained rather little. To continue, we need the notion of the *mean* of such a function f , and also some generalisation of the Fourier series expansions. For this, we have to be able to (locally) integrate the function f . A natural approach is provided by *Bochner's integral* [8], which can be viewed as an extension of the Lebesgue integral to functions with values in a general Banach space; see [13, App. E] as well as [40, Sec. V.5] for modern expositions.

With this extension, most of our previous results remain true, with the only change that the coefficients a_ℓ are now elements of \mathbb{X} rather than complex numbers. For instance, one has $M(Q_m) = a_0$ for the trigonometric polynomial of Eq. (5.1), and the analogue of Proposition 4.2 holds without change. Now, also the consecutive steps have their natural analogues, and we obtain the following result.

Theorem 5.2. *Let $(\mathbb{X}, |\cdot|)$ be a Banach space, and let $f: \mathbb{R} \rightarrow \mathbb{X}$ be Bohr-almost periodic. Then, for any fixed $\alpha \in \mathbb{R}$ with $|\alpha| > 1$, one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\alpha^n x) = M(f),$$

which holds for a.e. $x \in \mathbb{R}$. □

The extension to almost periodic functions in the Stepanov sense works in complete analogy, and we leave further steps in this direction to the reader.

ACKNOWLEDGEMENTS

MB would like to thank Jean-Pierre Conze, Michael Coons, Uwe Grimm and Nicolae Strungaru for discussions and helpful comments. Financial support by the German Research Council (DFG) through CRC 701 is gratefully acknowledged.

REFERENCES

- [1] Apostol T.M. *An Introduction to Analytic Number Theory* (Springer, New York, 1976).
- [2] Baake M. and Gähler F. Pair correlations of aperiodic inflation rules via renormalisation: Some interesting examples, *Topol. Appl.* **205** (2016), 4–27; [arXiv:1511.00885](#).
- [3] Baake M. and Grimm U. *Aperiodic Order. Vol. 1: A Mathematical Invitation* (Cambridge University Press, Cambridge, 2013).
- [4] Baake M. and Haynes A. A measure theoretic result for approximation by Delone sets, *Preprint* [arXiv:1702.04839](#).
- [5] Baake M. and Moody R.V. Weighted Dirac combs with pure point diffraction, *J. Reine und Angew. Math. (Crelle)* **573** (2004), 61–94; [arXiv:math.MG/0203030](#).
- [6] Besicovitch A.S. *Almost Periodic Functions*, reprint (Dover, New York, 1954).
- [7] Besicovitch A. and Bohr H. Some remarks on generalisations of almost periodic functions, *Dan. Math. Fys. Medd.* **8** (1927), 1–31.
- [8] Bochner S. Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind, *Fundamenta Math.* **20** (1933), 262–276.
- [9] Bohr H. *Almost Periodic Functions*, reprint (Chelsea, New York, 1947).
- [10] Bufetov A.I. and Solomyak B. On the modulus of continuity for spectral measures in substitution dynamics, *Adv. Math.* **260** (2014), 84–129; [arXiv:1305.7373](#).
- [11] Bugeaud Y. *Distribution Modulo One and Diophantine Approximation*, (Cambridge University Press, Cambridge, 2012).
- [12] Cigler J. Ziffernverteilung in ϑ -adischen Brüchen, *Math. Z.* **75** (1964), 8–13.
- [13] Cohn D.L. *Measure Theory*, 2nd ed. (Birkhäuser/Springer, New York, 2013).
- [14] Corduneanu C. *Almost Periodic Functions*, 2nd English ed. (Chelsea, New York, 1989).
- [15] Cornfeld I.P., Fomin S.V. and Sinai Ya.G. *Ergodic Theory*, SCSM 245 (Springer, New York, 1982).
- [16] de Bruijn N.G. and Post K.A. A remark on uniformly distributed sequences and Riemann integrability, *Nederl. Akad. Wetensch. Proc. Ser. A* **71** (1968), 149–150; now available as *Indag. Math. (Proc.)* **30** (1968), 149–150.
- [17] Drmota M. and Tichy R.F. *Sequences, Discrepancies and Applications*, LNM 1651 (Springer, Berlin, 1997).
- [18] Fan A.-H., Saussol B. and Schmeling J. Products of non-stationary random matrices and multiperiodic equations of several scaling factors, *Pacific J. Math.* **214** (2004), 31–54; [arXiv:math/0210347](#).
- [19] Hardy G.H. *Divergent Series* (Clarendon Press, Oxford, 1949).
- [20] Harman G. *Metric Number Theory* (Oxford University Press, New York, 1998).
- [21] Hartinger J., Kainhofer R.F. and Tichy R.F. Quasi-Monte Carlo algorithms for unbounded, weighted integration problems, *J. Complexity* **20** (2004), 654–668.
- [22] Kac M. On the distribution of values of sums of the type $\sum f(2^k t)$, *Ann. Math.* **47** (1946), 33–49.
- [23] Kamarul Haili H. and Nair R. The discrepancy of some real sequences, *Math. Scand.* **93** (2003), 268–274.
- [24] Katznelson Y. *An Introduction to Harmonic Analysis*, 3rd ed. (Cambridge University Press, Cambridge, 2004).
- [25] Kuipers L. and Niederreiter H. *Uniform Distribution of Sequences*, reprint (Dover, New York, 2006).
- [26] Leobacher G. and Pillichshammer F. *Introduction to Quasi-Monte Carlo Integration and Applications* (Birkhäuser, Basel, 2014).
- [27] Lesigne E. Loi des grands nombres pour des sommes de Riesz–Raikov multidimensionnelles, *Compositio Math.* **110** (1998), 39–49.

- [28] Loomis L.H. *Introduction to Abstract Harmonic Analysis*, reprint (Dover, New York, 2011).
- [29] Moody R.V. and Strungaru N. Almost periodic measures and their Fourier transforms. In *Aperiodic Order. Vol. 2: Crystallography and Almost Periodicity*, eds. Baake M. and Grimm U. (Cambridge University Press, Cambridge, 2017).
- [30] Parry W. On the β -expansion of real numbers, *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.
- [31] Petit B. Le théorème limite central pour des sommes de Riesz–Raikov, *Probab. Th. Rel. Fields* **93** (1992), 407–438.
- [32] Rényi A. Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hungar.* **8** (1957), 477–493.
- [33] Rio E. Lois fortes des grands nombres presque sûres pour les sommes de Riesz–Raikov, *Probab. Th. Rel. Fields* **118** (2000), 342–348.
- [34] Sobol I.M. Calculation of improper integrals using uniformly distributed sequences, *Soviet Math. Dokl.* **14** (1973), 734–738.
- [35] Stepanoff W. Über einige Verallgemeinerungen der fastperiodischen Funktionen, *Math. Ann.* **45** (1925), 473–498.
- [36] Strungaru N. Almost periodic pure point measures. In *Aperiodic Order. Vol. 2: Crystallography and Almost Periodicity*, eds. Baake M. and Grimm U. (Cambridge University Press, Cambridge, 2017).
- [37] Takahashi S. On the distribution of values of the type $\sum f(q^k t)$, *Tohoku Math. J.* **14** (1962), 233–243.
- [38] Walters P. *An Introduction to Ergodic Theory*, reprint (Springer, New York, 2000).
- [39] Weyl H. Über die Gleichverteilung von Zahlen mod. Eins, *Math. Annalen* **77** (1916), 313–352.
- [40] Yoshida K. *Functional Analysis*, SCSM 123, 6th ed. (Springer, Berlin, 1980).

FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT BIELEFELD,
 POSTFACH 100131, 33501 BIELEFELD, GERMANY
E-mail address: mbaake@math.uni-bielefeld.de

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HOUSTON,
 3551 CULLEN BLVD., HOUSTON, TX 77204-3008, USA
E-mail address: haynes@math.uh.edu

FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT JENA,
 ERNST-ABBE-PLATZ 2, 07743 JENA, GERMANY
E-mail address: daniel.lenz@uni-jena.de